Application/Control Number: 09/966,259   Docket No.: 2000-0572
Art Unit: 2655

REMARKS

Reconsideration and allowance are requested. Claims 1 – 32 are rejected and no

claims are amended.

Rejection of Claims 1, 5, 7, 8, 10, 12-13, 19, 21, 26 – 27 and 31-32 Under Section 102(e)

The Examiner rejects claims 1, 5, 7, 8, 10, 12-13, 19, 21, 26 – 27 and 31-32 under

Section 102(e) as being anticipated by U.S. Pub. No. 2002/0052742 to Thrasher et al.

("Thrasher et al."). Applicants traverse this rejection and submit that Thrasher et al fail to

teach each claim limitation.

We first turn to independent claims 1, 13 and 21. The Examiner rejects these claims

asserting that the language model taught by Thrasher et al. in FIG. 2, feature 110 and

paragraph 0038 anticipates the controller (of claim 1) that determines the data of the at least

one communications device, transducer, vocal information and acoustic environmental data

and then compensates at least one speech recognition model to reflect the data. Applicants

traverse this rejection and submit that it is clear that Thrasher et al. do not teach a controller

that compensates at least one speech recognition model. There is simply no compensation of

the at least one speech recognition model that occurs in the teachings of Thrasher et al.

For example, paragraph 0038 of Thrasher et al. discusses a speech recognition engine

103 that generates the N-best alternatives for recognized speech. They explain that where the

speech recognizer incurs errors in speech recognition, the user can provide input to signal the

N-best alternative generator to access the reference path 117 (this is the recognized speech),

the hypothesis lattice 116 and the language model 110. The N-best alternative generator

utilizes these inputs, as well as user input associated with the text to be corrected, and

generates the top N-best alternatives to the word or phrase selected by the user. In paragraph

0032, Thrasher et al. teach about the language model 110 and how it provides a set of

likelihoods that a particular word will appear in a particular language. Importantly, Thrasher

et al. teach that the process of speech recognition utilizes the language model to identify a

2

most likely sequence of words from the speech. Paragraph 0038 deals with a way to correct

for incorrect recognitions wherein the N-best alternatives generator accesses the language

model 110 as well as other inputs to generate the N-best alternatives. There is simply no

teaching in Thrasher et al. regarding compensating the at least one speech recognition model.

The language model in Thrasher et al. does not change but is merely accessed once during

original speech recognition and again in the process of generating the N-best alternatives list

in the correction process of paragraph 0038.

The Examiner, on page 3 of the current office action, also makes no argument that the

language model is compensated. He sets for the same analysis above regarding how Thrasher

et al. handle incorrect speech recognition by accessing for a second time the *same* language

model as part of generating the N-best alternatives list. Accessing the same language model

several times is different from compensating a speech recognition model as in the claims.

Accordingly, since claims 1, 13 and 21 require compensation of the speech

recognition model to reflect the data, and where Thrasher et al. fail to teach any type of

change (compensation) to the language model based on such data, Applicants submit that

these independent claims are patentable and in condition for allowance.

Rejection of Claims 2-4, 6, 9, 11, 18, 20, 22-24, 25 and 28 Under Section 103(a)

The Examiner rejects claims 2-4, 6, 9, 11, 18, 20, 22-24, 25 and 28 under Section

103(a) as being unpatentable over Thrasher et al. in view of U.S. Pat. No. 6,304,844 to Pan et

al. ("Pan et al."). Applicants traverse this rejection because Applicants submit that there is no

motivation or suggestion to combine these references to reject the claims.

As noted in the previous office action, Applicants submit that there is no motivation

or suggestion to combine Thrasher et al. with Pan et al. In response, on page 3 of the current

Office Action, the Examiner asserts that "each claim limitation teaches from the reference the

motivation to combine." Applicants respectfully submit that this analysis represents an

erroneous application of the doctrine of obviousness. The source for the motivation to

3

Application/Control Number: 09/966,259   Docket No.: 2000-0572
Art Unit: 2655

combine two references *cannot* come from Applicants' claim limitations. MPEP 2143.01

sets forth the source for the motivation to combine: a teaching, suggestion or motivation

found (1) either in the references themselves; or (2) in the knowledge generally available to

one of ordinary skill in the art. Since the patent application is kept secret upon filing for at

least 18 months, the claim limitations cannot be the source of motivation for combining

references. Further, as set forth in MPEP 2143.01, the test for obviousness is what the

combined teachings of the references would have suggested. All the teachings in the prior art

must be considered. Where references teachings conflict as in Thrasher et al. (full blown

word/phrase recognition system on a powerful computer system) and Pan et al. (letter

recognition system for low-power, low memory portable devices), the Examiner must weight

the power of each reference to suggest solutions and consider the degree to which one

reference might accurately discredit or avoid suggesting combining with the other.

Furthermore, the Examiner noted on page 3 that the page and line number in which

information from each reference can be found followed the analysis. Simply pointing out

some specific language in each reference is insufficient to establish a prima facie case of

obviousness. Applicants provided in the previous Response an analysis that considered all

the teachings of the prior art and examined and weighed the suggestive power of each

reference (not just a cited portion of the reference) to be combined with the other. The MPEP

also states that the mere fact that references <u>can</u> be combined or modified does not render the

resultant combination obvious unless the prior art suggests the desirability of the

combination. MPEP 2143.01 (emphases in original). Applicants respectfully submit that the

Examiner is operating within this particular realm - wherein the references perhaps <u>can</u> be

combined but where the references do not contain the requisite suggestion or motivation to be

combined.

Applicants remind the Examiner that the standard by which obviousness is judged is

only by a preponderance of the evidence. MPEP 2142. As previously explained with regards

4

Application/Control Number: 09/966,259   Docket No.: 2000-0572
Art Unit: 2655

to the overall teachings of the prior art references and how they teach away from any

suggestion or motivation to combine, Applicants submit that the preponderance of the

evidence leans away from a conclusion of obviousness to combine.

Further reasons do exist as well. For example, with regards to claims 2 and 22, the

Examiner asserts that Pan et al. discloses "a distortion value (distortion scores) related to a

transducer of a mobile communications device (column 6, lines 25 - 57)...." Applicants

submit that Pan et al. fail to teach that the distortion score relates to anything specific and

there is no mention of the transducer. The "score" taught by Pan et al. is disclosed as the

closeness of each inputted combination to a word in the vocabulary database. The

"closeness" index is based upon the cepstral distance between the input waveform and the

stored vocabulary waveforms, thereby generating "distortion scores". There are a number of

possible reasons for the distortion score. For example, the transducer may be the same for

recording the vocabulary database and the input speech but the speaker has a cold when

inputting the speech. Other environmental factors such as wind or background noise may

cause changes in the input speech relative to the vocabulary database and so forth. The

Examiner simply assumes that the distortion score is related to the transducer. This

connection between a distortion score and the transducer is not stated or suggested by the

reference. Furthermore, the claims require a distortion value related to a transducer and Pan

et al. simply does not teach anything in this cited portion about a transducer. Therefore, even

if combined, these references fail to teach the limitations of claims 2 and 22.

Regarding claims 3 and 23, Pan et al. fail to teach the limitation recited in these

claims. The Examiner asserts that Pan et al. teach an acoustic environmental data and a

background noise value that corresponds to an operating environment of the device.

Applicants traverse this analysis. Pan et al. mention background noise in column 10, line 16

but fail to teach a background noise value. Pan et al. teach away from arriving at a

background noise value by the nature of their invention. The Pan et al. invention deals with

5

Application/Control Number: 09/966,259   Docket No.: 2000-0572
Art Unit: 2655

the problems of background noise in speech recognition by increasing accuracy of

recognition through receiving spelling for input utterances. Speaking letters instead of words

increases the use of typical pauses between letters and thus reduces the effect of background

noise. Thus, rather than teach about a background noise value, Pan et al. simply deal with

background noise by having the user input letters which are easier to recognize.

Therefore, even if combined, Applicants submit that claims 3 and 23 are patentable

because the combination of references fails to teach each claim limitation.

Regarding claims 9 and 28, Applicants respectfully traverse the Examiner's equating

the limitation that the controller is a network server with the teachings in Pan et al., figure 1,

element 103. Element 103 in figure 1 of Pan et al. is a pronunciation database. Pan et al.

teach away from their invention being a "network server" and clearly teach a mobile, low

power consumption and low computational power. Col. 2, lines 54 - 65. To equate Pan et

al.'s figure 1 with a network server cuts directly against the express teachings of Pan et al.

that their invention relates to a device that has low processing power and limited memory

storage. Such a computing device is certainly not a network server. Therefore, even if

combined, these references teach away from the limitations of claims 9 and 28.

Regarding claim 11, Pan et al. fail to teach the limitations of this claim. Column 4,

lines 1 - 4 of Pan et al. teach storing a user's pronunciation of letters in a pronunciation

database. There is no mention of "administrative information relating to an end user" as is

recited in the claim. Furthermore, column 10, lines 9 - 27 fail to mention anything regarding

a probability value that represents a probability that the end user will be in a particular

background environment. The only thing mentioned in this portion of Pan et al. is that there

is background noise that makes accurate endpoint detection difficult. As mentioned above,

Pan et al. solve this problem by making the user's speech related to letters rather than words.

The limitation in claim 11 regarding a probability value is simply not mentioned by Pan et al.

Therefore, even if combined, these references fail to teach the limitations of claim 11.

Application/Control Number: 09/966,259   Docket No.: 2000-0572
Art Unit: 2655

Regarding claim 18, Pan et al. fail to mention that the vocal information represents a variability that exists in vocal tract shapes among speakers of a group. Pan et al. do mention taking into consideration vocal tract characteristics, but they are only mentioned with reference to individual speakers. There is no mention of variability and no mention of variability of speakers within a group. Therefore, even if combined, the references fail to teach each limitation of claim 18.

Regarding claim 20, Applicants submit that column 5, lines 5 - 8 of Pan et al., which mentions storing pronunciations of a user in a pronunciations database, simply does not equate to the limitation of claim 1 which involves storing personal account information of the user. There is a difference between the actual pronunciation data that is taught as being stored by Pan et al. and a user's personal account information for each end user which is recited in claim 20. Therefore, the limitations of claim 20 are not taught by the references even if combined.

Regarding claim 24, the Examiner cites column 10 and column 2 as teaching the limitation where the data is received from a cellular telephone. The only mention in these portions of Pan et al. with regards to a cellular telephone is that a phone may be used for speech recognition. There is no teaching that a cellular phone transmits at least one of data associated with a communications device, transducer, vocal information and acoustic environment data to an automatic speech recognition system. Therefore, Applicants submit that claim 24 is patentable even if the references are combined.

Each claim that depends from claims 1, 13 or 21 are patentable inasmuch as the parent claims are patentable.

Rejection of Claim 14 Under 35 U.S.C. Section 103(a)

The Examiner rejects claim 14 under 35 U.S.C. Section 103(a) as being unpatentable over Thrasher et al. in view of U.S. Pat. No. 6,720,888 to Eagleson et al. ("Eagleson et al."). The Examiner did not make any comments in response to Applicants arguments against the

7

Application/Control Number: 09/966,259   Docket No.: 2000-0572
Art Unit: 2655

combination of Eagleson et al. with Thrasher et al. In sum, Eagleson et al. has nothing to do

with speech recognition technology. There is no reason to combine the correction method for

user-selected recognized speech in Thrasher et al. with the signpost mobile device tracking

invention of Eagleson et al. The technology subject matter in these two references is entirely

different and one of skill in the art would not find any motivation to combine their teachings.

Therefore, Applicants submit that claim 14 is patentable over these references.

Rejection of Claims 15-16 Under 35 U.S.C. Section 103(a)

The Examiner rejects claims 15 and 16 under 35 U.S.C. Section 103(a) as being

unpatentable over Thrasher in view of U.S. Patent No. 6,219,645 ("Byers). Claims 15 and 16

depend from claim 13, which as discussed above, is patentable over Thrasher et al. because

that reference fails to teach a second section that compensates a speech recognition model.

The language model in Thrasher et al. is accessed several times but never compensated based

on data as is recited in the claim. Therefore, these dependent claims are patentable.

Rejection of Claim 17 Under 35 U.S.C. Section 103(a)

The Examiner rejects claim 17 under 35 U.S.C. Section 103(a) as being unpatentable

over Thrasher et al. in view of U.S. Patent No. 5,778,336 ("Chou"). Applicants traverse this

rejection and submit that there is no suggestion or motivation to combine Thrasher et al. with

Chou. Chou teaches a speech coding and joint data/channel bias estimation. The context of

the application of the Chou invention, as stated on column 2, line 66-67 is in a

communication system where data transmission is modeled by a dispersive channel with

additive noise. In a communications channel such as a wireless channel, the speech signal is

degraded due to environmental and channel distortion. Col. 1, lines 12 - 35. Where Chou

focuses on a communication system channel with channel distortion, Thrasher et al. teach a

system wherein the microphone is connected directly to the computer having the speech

recognition system. See FIGs. 1 and 2 of Thrasher et al. and associated description. Thrasher

et al.'s invention is not taught as in the context of a communications system wherein the

8

Application/Control Number: 09/966,259   Docket No.: 2000-0572
Art Unit: 2655

channel distortion exists. Where the microphone is connected directly to the computer, there

is no channel distortion that is introduced by way of environmental conditions or the type of

distortion that arises from a wireless channel. Therefore, one of skill in the art (by a

preponderance of the evidence) would not likely combine these references because there is no

need for correcting for channel distortion in Thrasher et al. The Examiner states that the

motivation to combine is that it would be obvious to modify Thrasher et al.'s system with

Chou's distortion value based on a difference between an actual transducer and a response

characteristic of a transducer to train the speech recognition model for the purpose of

minimizing the adverse effects of acoustic mismatch (referencing col. 1, lines 13 - 32 of

Chou). However, the very reason cited by the Examiner from Chou, the issue of acoustic

mismatch, is nonexistent in Thrasher et al. because there is no "communications channel"

that introduces environment distortion as explained above.

Accordingly, one of skill in the art, rather than find motivation to combine Chou with

Thrasher et al., would recognize that Chou's teachings regarding distortion in communication

channels are inapplicable to Thrasher et al. because there is no channel having the distortion

problems introduced in Chou. For this reason, Applicants submit that claim 17 is patentable

and in condition for allowance.

Rejection of Claim 29 Under 35 U.S.C. Section 103(a)

The Examiner has rejected claim 29 under 35 U.S.C. Section 103(a) as being

unpatentable over Thrasher in view of Pan et al. and further in view of Byers. Applicants

submit that claim 29 depends from claims 23 and 21, which as discussed above, are

patentable. Therefore, this claim is patentable as well. An additional reason against any

motivation to combine these references is that Byers focuses on the plurality of speech

recognition systems working within the volume of a room. Pan et al. as discussed above is a

mobile device application that teaches away from using a server or more powerful computer

to handle speech recognition. Byers's invention clearly is not a mobile device but is set forth

9

Application/Control Number: 09/966,259   Docket No.: 2000-0572
Art Unit: 2655

in a room with multiple microphones and a directional ASR controller. There is no mobility

suggested in Byers invention. Therefore, Applicants submit that there is no reason or

motivation to combine Byers with Pan et al.

Rejection of Claim 30 Under 35 U.S.C. Section 103(a)

The Examiner has rejected claim 30 under 35 U.S.C. Section 103(a) as being

unpatentable over Thrasher in view of Pan et al. and further in view of Chou. As discussed

above, there is no motivation or suggestion to combine Thrasher et al. with Chou. Therefore,

for this reasons as well as that claim 30 depends from an allowable claim, Applicants submit

that claim 30 is patentable.

CONCLUSION

Having addressed the rejection of claims 1 - 32, Applicants respectfully submit that

the subject application is in condition for allowance and a Notice to that effect is earnestly

solicited.

Respectfully submitted,

Date: July 6, 2005   By: /Thomas M. Isaacson/

Correspondence Address: Thomas M. Isaacson
Samuel H. Dworetsky   Attorney for Applicants
AT&T Corp.   Reg. No. 44,166
Room 2A-207   Phone: 410-414-3056
One AT&T Way   Fax No.: 410-510-1433
Bedminster, NJ 07921

10